

Method and decoder for composing a scene**FIELD OF THE INVENTION**

5 The present invention relates to a method for composing a scene containing a plurality of objects, an object comprising chrominance and luminance components, a chrominance value being associated with a set of at least two luminance values, said method comprising a step of blending a first object with a second object resulting in a blended object.

10 The invention also relates to a decoder, said decoder implementing said method. Such a method may be used in, for example, a MPEG-4 video communication system.

BACKGROUND OF THE INVENTION

A visual scene can be composed of objects of different type such as natural video, two-dimensional, three-dimensional synthetic images, graphics or text.

15 A video communication system, like for example a mobile telecommunication system, typically comprises a first mobile phone and a second mobile phone, including an encoder and a decoder, respectively. In order to transmit a visual scene from the first mobile phone to the second mobile phone, the different objects are encoded by the first mobile phone and transmitted independently to the second mobile phone. At the decoder side the scene is
20 recomposed by decoding the different objects and then by rendering the decoded objects. The rendering usually comprises a geometric transformation step able to map the object on the scene and a step of blending the mapped objects together, resulting in a final scene.

25 An object typically comprises chrominance U and V, and luminance Y components forming the color of the object. During the blending step of a current object N with an object M, a transparency value Nalpha can be associated to the object N, said transparency value influencing the chrominance and luminance values Pyuv of a blending object P such that:

$$Pyuv = [(255 - Nalpha) * Myuv + Nalpha_y * Nyuv] / 255$$

30 The transparency of a pixel of a blended object is obtained according to the shape of said pixel. The transparency permits to determine which object is behind or in front of another object in the scene. The shape is associated with a pixel of a frame of a scene. It permits to know if a pixel is outside or inside an object. If it is outside, the shape has a null value, otherwise it has a value equal to 1.

For a pixel having coordinates (x,y), the shape value 0 or 1 corresponding to these coordinates is taken and attributed to the luminance of this pixel. In the case where the

luminance values the crosses), the shape used for the chrominance value U or $V(x,y)$ is the shape of the pixel of coordinates $(2x-1, 2y-1)$, i.e. the top right pixel of a square of 2×2 pixels associated with the chrominance value.

One inconvenient of this solution is that the chrominance of the blended object is false. This results in a bad visual quality for objects having an arbitrary shape. This observation is illustrated in the example of an object constituted of the letter "e" and represented by stripes in Fig. 2a to 2c. Fig. 2a is the original object, Fig. 2c is the result of the composition and Fig. 2b explains how the result is obtained. Each square represents a pixel. If one takes the pixel whose coordinates are $(4,2)$, the pixel whose shape will be used for the chrominance component is the pixel $(3,1)$ as shown in Fig. 2b. The value of this shape is 0, that is to say it is outside the letter "e", and the chrominance value will not be taken into account, as shown in Fig. 2c where the pixel value is represented by dots. The result is false because the chrominance value should have been taken into account and the result should have been the same as in Fig. 2a.

SUMMARY OF THE INVENTION

Accordingly, it is an object of the invention to provide a method and a decoder for composing a scene containing a plurality of objects, an object comprising chrominance and luminance components, a chrominance value being associated with a set of at least two luminance values, said method comprising a step of blending a first object with a second object resulting in a blended object, which gives a better visual quality of said blended object.

To this end, there is provided a method wherein the blending step comprising the sub-steps of:

- generating a luminance component of the blended object from the corresponding luminance components of the first and second objects and from a first composition function, and
- generating a chrominance component of the blended object from the corresponding chrominance components of the first and second object and from a second composition function, the second composition function depending on a set of associated values of the first composition function.

In addition, there is provided a decoder comprising blending means, said blending means comprising:

- luminance generation means for generating a luminance component of the blended object from the corresponding luminance components of the first and second objects and from a first composition function, and
- chrominance generation means for a generating chrominance component of the
5 blended object from the corresponding chrominance components of the first and second object and from a second composition function, the second composition function depending on a set of associated values of the first composition function.

As we will see in detail further on, using a set of associated values of the first composition function for the second composition function permits to have a better precision
10 for the chrominance component of an object.

According to a first non-limited embodiment, the first composition function depends on a shape component.

In a non-limited variant of this first non-limited embodiment, a chrominance value is associated with 4 luminance values and 4 shape values, the second composition function
15 being an 'OR' function between the 4 associated shape values.

According to a second non-limited embodiment, the first composition function is based on a transparency component.

In a non-limited variant of this second non-limited embodiment a chrominance value is associated with 4 luminance values and 4 transparency values, the second composition
20 function being an average of the 4 transparency values.

BRIEF DESCRIPTION OF THE DRAWINGS

Additional objects, features and advantages of the invention will become apparent upon reading the following detailed description and upon reference to the accompanying
25 drawings in which:

- Fig. 1 illustrates a distribution of the chrominance and luminance components in the case of the 4:2:0 format,
- Fig. 2a, 2b and 2c illustrate a result of a chrominance computation of a blended object of a scene according to the prior art, and
- 30 - Fig. 3 is a block diagram of a decoder.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a method for composing a scene, said scene containing a plurality of objects, an object comprising chrominance and luminance

components, a chrominance value being associated with a set of at least two luminance values.

In the following description, the invention is depicted in the case of a video format 4:2:0, i.e. one chrominance U or V value corresponds to four luminance Y values, the chrominance being sub-sampled in the horizontal and vertical directions, but it will be apparent to a person skilled in the art that the invention is applicable to any video format in which the chrominance component is sub-sampled compared to the luminance component.

In the following description, well-known functions or constructions by a person skilled in the art are not described in detail since they would obscure the invention in unnecessary detail.

Fig.3 represents a general overview of a decoder structure. The same decoding scheme is applied when decoding all the objects of a given session.

The decoder comprises a demultiplexer DEMUX (31), a shape decoder SHDEC (32), a motion decoder MDEC (33) and a texture decoder YUVDEC (34) for decoding the luminance and chrominance components. The reconstructed object is obtained by the right combination of the shape, texture and motion information thanks to a motion compensation circuit MC (35), a reconstruction circuit (36) and a picture memory MEM (37), according to a principle known to a person skilled in the art.

After the decoding of the different objects of a given session, an inverse affine transform, such as rotation, a zoom or a reduction for example, is applied on each object of a scene. To this end, a well-known technique called « backward mapping » is used. The inverse affine transform is applied on the coordinates (x,y) of each pixel of said image. The luminance, chrominance and transparency components of a pixel of the current object are then obtained by an interpolation, a bilinear interpolation for example, between four pixels that are surrounding the inverse transformed pixel having coordinates (X,Y).

After the inverse affine transform, the blending step also called composition is performed. The objects are recursively blended in a predetermined order.

In this step, a transparency value is associated to each object of the scene. Practically, it permits to see which object is behind or in front of another object. Each object has its own luminance Y, chrominance U and V, and transparency alpha values. Blending is done sequentially. Usually, the objects are composed two by two beginning by those on the backward of a scene. For example, if a first object N is over-layered over a second object M

to generate a blended object P. The YUV pixel values of the blended object P depend on those of the objects M and N and on the transparency value of the current object N. Hence:

$$Py = [(255 - Nalpha_y) * My + Nalpha_y * Ny] / 255$$

$$Puv = [(255 - Nalpha_uv) * Muv + Nalpha_uv * Nuv] / 255$$

5 where:

- Py and Puv are the luminance Y and chrominance UV components, respectively, of the blended object P,
- My and Muv are the luminance Y and chrominance UV components, respectively, of the second object M, which is for example the background of the scene, and which has been computed in a preceding blending step and memorized,
- 10 - Ny, Nuv are the luminance Y and chrominance UV components, respectively, of the first object N being processed,
- Nalpha_y and Nalpha_uv are the transparency components applied to the pixels of an object N in order to obtain the Py and Puv components, respectively,
- 15 - the luminance value is comprised between 0 and 255.

A transparency component Nalpha is the product of three terms:

- a global alpha value, which is applied to all the pixels of an object,
- some gray-alpha values, one value being associated with one pixel,
- some binary shape values, one value being associated with one pixel.

20 Note that the transparency and the shape components have the same resolution as the luminance.

More particularly, the luminance component of a blended object P is generated from the corresponding luminance components of the first and second objects N and M and from a first composition function. The chrominance component of the blended object P is generated from the corresponding chrominance components of the first and second objects N and M and from a second composition function.

Thus, the transparency Nalpha_y of the luminance component is function of the first composition function, and the transparency Nalpha_uv of the chrominance component is function of the second composition function. Advantageously, the second composition function depends on a set of associated values of the first composition function. It permits to be more precise in the computation of the chrominance component Puv.

According to a first non-limited embodiment of the invention, the first composition function is based on a binary shape component and the transparency component $N_{\alpha_{uv}}$ is the product of global-alpha value, gray-alpha value and a shape value $Shape_{uv}(x,y)$.

In this case, the second composition function is an OR of the shape of the
5 corresponding four pixels such that:

$Shape_{uv}(x,y) = Shape_y(2x-1,2y-1) \text{ OR } Shape_y(2x,2y-1) \text{ OR } Shape_y(2x-1,2y) \text{ OR } Shape_y(2x,2y)$ if pixels are numbered starting from 1.

For example for the current pixel whose coordinates are $(x,y) = (2,1)$, the second composition function is:

10 $Shape_{uv}(2,1) = shape_y(3,1) \text{ OR } shape_y(4,1) \text{ OR } shape_y(3,2) \text{ OR } shape_y(4,2)$,
the different shapes values $Shape_y()$ taking the values 0 or 1. Hence, the chrominance value is taken into account if at least one luminance value is taken into account, i.e. belongs to the object N.

With this second composition function, the resulting blended object "e" of Fig. 2c has
15 a much better quality, as it is the original "e".

As for the gray-alpha value of the transparency $N_{\alpha_{uv}}$, one can take either the current pixel gray-alpha value:

$gray_alpha_{uv}(x,y) = gray_alpha_y(2x-1,2y-1)$,

or one can take the gray-alpha value of the four associated pixel, such that:

20 $gray_alpha_{uv}(x,y) = [gray_alpha_y(2x-1,2y-1) + gray_alpha_y(2x,2y-1) + gray_alpha_y(2x-1,2y) + gray_alpha_y(2x,2y)]/4$.

According to a second non-limited embodiment of the invention, the first composition function is based on the transparency component.

25 In this case, the second composition function is an average of the transparency of four associated pixels such that:

$N_{\alpha_{uv}}(x,y) = [N_{\alpha_y}(2x-1,2y-1) + N_{\alpha_y}(2x,2y-1) + N_{\alpha_y}(2x-1,2y) + N_{\alpha_y}(2x,2y)]/4$

It is to be noted that the average can be computed in a different way, for example
30 using weighting coefficient.

As the value of the transparency is comprised between 0 and 255, the result gives a better precision than if only the shape is used as in the first embodiment.

It is to be understood that the present invention is not limited to the aforementioned embodiments and variations and modifications may be made without departing from the spirit and scope of the invention as defined in the appended claims. In the respect, the following closing remarks are made.

5 It is to be understood that the present invention is not limited to the aforementioned video application on mobile phone. For example it can be applied on a television whenever there are some text message and video.

It is to be understood that the method according to the present invention is not limited to the aforementioned implementation.

10 There are numerous ways of implementing functions of the method according to the invention by means of items of hardware or software, or both, provided that a single item of hardware or software can carries out several functions. It does not exclude that an assembly of items of hardware or software or both carry out a function, thus forming a single function without modifying the method for composing a scene in accordance with the invention.

15 Said hardware or software items can be implemented in several manners, such as by means of wired electronic circuits or by means of an integrated circuit that is suitable programmed respectively. The integrated circuit can be contained in a computer or in an encoder. In the second case, the decoder comprises luminance generation means for generating a luminance component of the blended object from the corresponding luminance
20 components of the first and second objects and from a first composition function, and chrominance generation means for a generating chrominance component of the blended object from the corresponding chrominance components of the first and second object and from a second composition function, the second composition function depending on a set of values of the first composition function, as described previously, said means being hardware
25 or software items as above stated.

The integrated circuit comprises a set of instructions. Thus, said set of instructions contained, for example, in a computer programming memory or in a decoder memory may cause the computer or the decoder to carry out the different steps of the composing method.

30 The set of instructions may be loaded into the programming memory by reading a data carrier such as, for example, a disk. A service provider can also make the set of instructions available via a communication network such as, for example, the Internet.

Any reference sign in the following claims should not be construed as limiting the claim. It will be obvious that the use of the verb "to comprise" and its conjugations do not

exclude the presence of any other steps or elements besides those defined in any claim. The article “a” or “an” preceding an element or step does not exclude the presence of a plurality of such elements or steps.